**Problem Statement: -**

In this time and age of widespread internet usage, effective and targeted marketing play a vital role, a marketing company would like to develop their strategy by analyzing their customer data and how effectively they can do targeted marketing, for this historical data has been collected of users clicking on ad given different factors such as age, location, time of activity and more. Perform Logistic Regression on the given data and classify the user who click’s on ad’s and who does not click on ad.

Solution:

There are five basic steps when you’re implementing logistic regression:

1. Import the packages and classes you need.
2. Provide data to work with and eventually do appropriate transformations.
3. Create a regression model and fit it with existing data.
4. Check for collinearity between independent variables and between independent and dependent variables .
5. Check for overfitting issue
6. Apply the model for predictions.

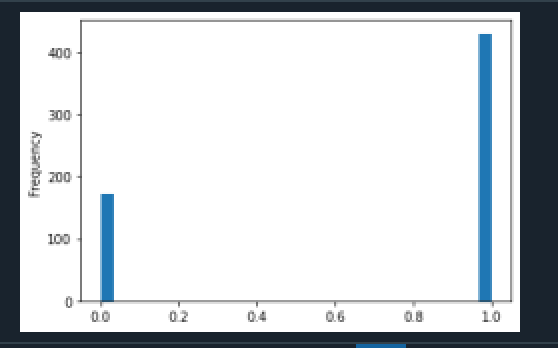
The first step is to import the package numpy and

importnumpyasnp

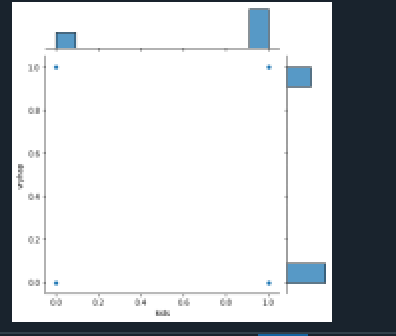
Now, you have all the functionalities you need to implement logistic regression.

relavancy check of dependent variable with independent variables by multiple plots

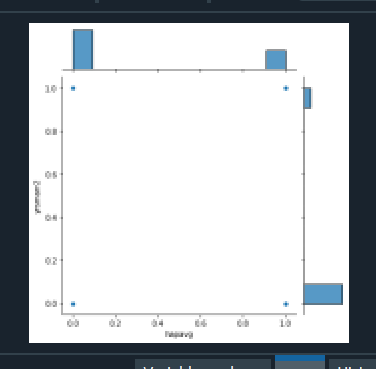
#data['kids'].plot.hist(bins=30)



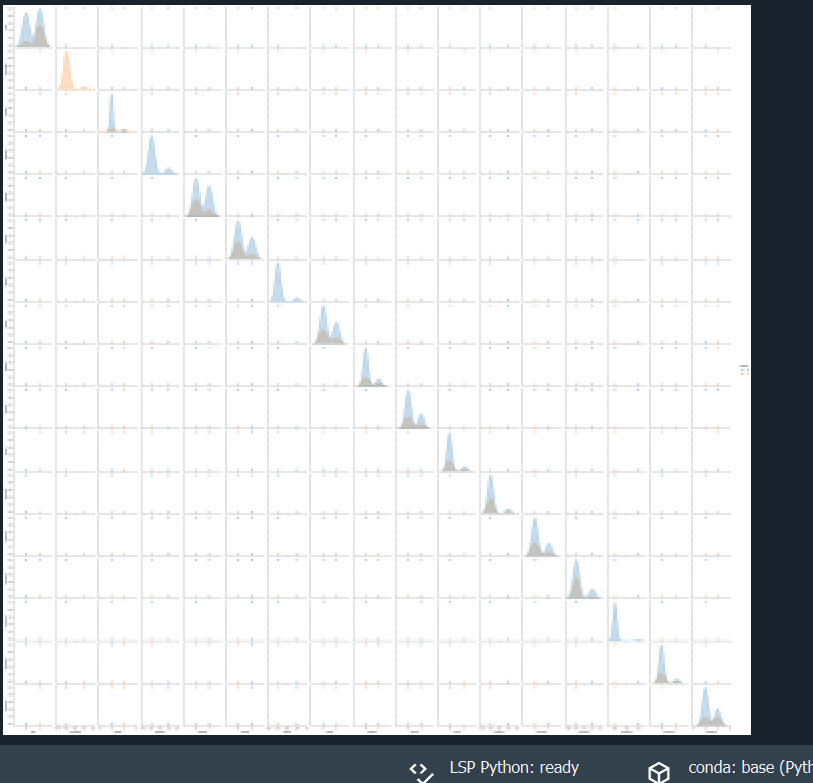
#sns.jointplot(x='kids',y='vryhap',data=data)



Let’s also determine the inter-dependency of ‘hapavg’ and ‘yrsmarr2’



Pairplot of data



Building the Model

Train and Test Data Sets

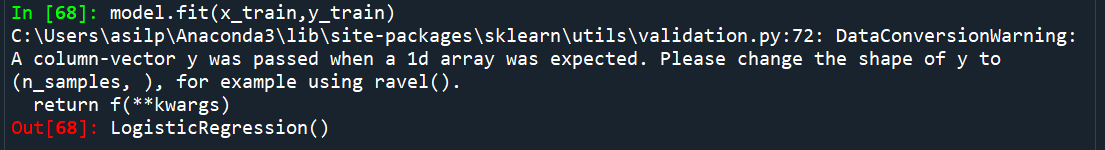
We must now divide our dataset into two distinct parts. A training set and a testing set. We can use scikit-learn for this.

Prior to segmenting the data, we will need to define our input and output variables i.e. X and y. For the purpose of this article, we will stick to numerical data when defining the input values for X.

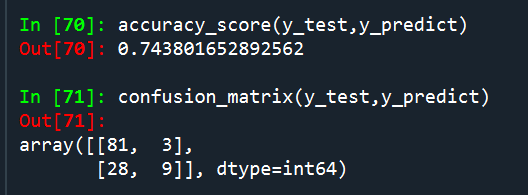


Training and testing the model

We will go ahead and import the Logistic Regression from the sklearn.linear\_model library:



By importing scikit-learn's ‘accuracy\_score’ we can swiftly assess the accuracy of our model and using confusion matrix we can visualize the accuracy of model



# Conclusion

The obtained result shows the power of Machine Learning algorithms. A basic Logistic Regression model has proven to be very successful in solving this classification problem.